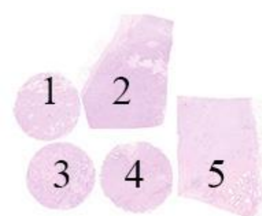
Assessment Run B32 2021

HER2 IHC

Purpose
Evaluation of the analytical accuracy of HER2 IHC tests performed by the NordiQC participants for demonstration and establishment of the HER2 protein expression level in breast carcinomas.
The HER2 IHC assays PATHWAY® (Ventana/Roche) and HercepTest™ (Dako/Agilent) were used as reference standard methods, and accuracy was evaluated in five breast carcinomas with the dynamic and critical relevant expression levels of HER2. The obtained score in NordiQC is indicative of the performance of the IHC tests used by the participants, but due to the limited number and composition of samples, internal validation and extended quality control, e.g. regularly measuring the HER2 results, is necessary and recommended.

Material
The slide to be stained for HER2 comprised the following 5 materials:

	IHC: HER2 Score* (0, 1+, 2+, 3+)	**FISH: HER2 gene/chr 17 ratio****
1. Breast carcinoma, no. 1	3+	4.6 (clusters) (amplified)
2. Breast carcinoma, no. 2	1-2+	1.0 (unamplified)
3. Breast carcinoma, no. 3	0-1+	1.1 (unamplified
4. Breast carcinoma, no. 4	3+	6.7 (clusters) (amplified)
5. Breast carcinoma, no. 5	2+	3.6 (amplified)



* HER2 immunohistochemical score (see table below) as achieved by using the two FDA / CE-IVD approved HER2 IHC assays, HercepTest™ (SK001, Dako/Agilent) and PATHWAY® (790-2991, Ventana/Roche), in NordiQC reference laboratories.
** HER2 gene/chromosome 17 ratio achieved using Zyto*Light* ® SPEC HER2/CEN 17 Dual Color FISH (Zytovision) in NordiQC reference laboratory.

All carcinomas were fixed for 24-48 h in 10% neutral buffered formalin.

IHC scoring system according to the 2018 ASCO/CAP guidelines:

Score 0	No staining is observed **or** membrane staining that is incomplete and is faint/barely perceptible and in ≤10% of tumor cells.
Score 1+	Incomplete membrane staining that is faint/barely perceptible and in >10% of tumor cells.
Score 2+	Weak to moderate complete membrane staining observed in >10% of tumor cells.
Score 3+	Circumferential membrane staining that is complete, intense, and in >10% of tumor cells*.

**Readily appreciated using a low-power objective and observed within a homogeneous and contiguous invasive cell population.*

Criteria for assessing a HER2 staining as **optimal** were:

- Staining corresponding to score 0 or 1+ in carcinoma no. 3.
- Staining corresponding to score 0, 1+ or 2+ in carcinoma no. 2.
- Staining corresponding to score 2+ or 3+ in carcinoma no. 5.
- Staining corresponding to score 3+ in carcinoma no. 1 and 4.
- No or only weak cytoplasmic reaction that did not interfere with the interpretation.

Staining was assessed as **good**, if (1) the HER2 gene amplified tumours no. 1 and 4 showed a 2+ reaction and the other breast carcinomas showed reaction pattern as described above (equivocal 2+ IHC staining should always be analyzed by ISH according to the ASCO/CAP guidelines) ***or*** (2) a less distinct and/or reduced number of neoplastic cells were demonstrated in the HER2 2+ gene amplified tumour no. 5 compared to the NordiQC reference standards determined by HercepTest™ and PATHWAY® ***or*** (3) a 2+ reaction was seen in the HER2 gene unamplified 0/1+ tumour no. 3.

Staining was assessed as **borderline,** if the signal-to-noise ratio was low, e.g., because of moderate cytoplasmic reaction, excessive counterstaining or impaired morphology hampering the interpretation.

Staining was assessed as **poor** in case of a false negative staining (e.g., the IHC 3+ tumours or the 2+ tumour with HER2 gene amplification showing a 0 or 1+ reaction) ***or*** a false positive staining (e.g. the IHC 2+ tumour without HER2 gene amplification showing a 3+ reaction).

Participation

Number of laboratories registered for HER2, run B32	402
Number of laboratories returning slides	364 (91%)

Results

At the date of assessment, 91% of the participants had returned the circulated NordiQC slides. All slides returned after the assessment were assessed and laboratories received advice if the result was insufficient, but the data were not included in this report.

In total 364 laboratories participated in this assessment. One submitted a slide stained for PR and was excluded from the data analysis. Of the remaining 363 participants, 82% achieved a sufficient mark (optimal or good).

The overall pass rate was slightly reduced compared to the level seen in the three latest assessment Runs B29-B31. Same assessment criteria have been applied and the difference might be correlated to more challenging material circulated for the run compared to previous runs. The pass rate, however, still at a satisfactory level.
In this assessment, the two established FDA-/CE-IVD approved HER2 IHC assays from Ventana/Roche, PATHWAY® 790-2991 and HER2/4B5 790-4493 provided the highest proportion of optimal results and a pass rate of 93% (all protocol settings).
The recently launched HercepTest™ GE001, Dako/Agilent, provided a high pass rate of 100% (all protocol settings) and 50% optimal results.
For the classical HercepTest™ (SK001, Dako/Agilent), Oracle™ (Leica Biosystems) and LDTs for HER2 IHC a significantly reduced pass rate was observed as illustrated in Graph 1, see below.

Assessment marks for IHC HER2 assays and HER2 antibodies are summarized in Table 1 (see page 3).

Graph 1. **Pass rates of the HER2 IHC assessments in the NordiQC breast cancer module**

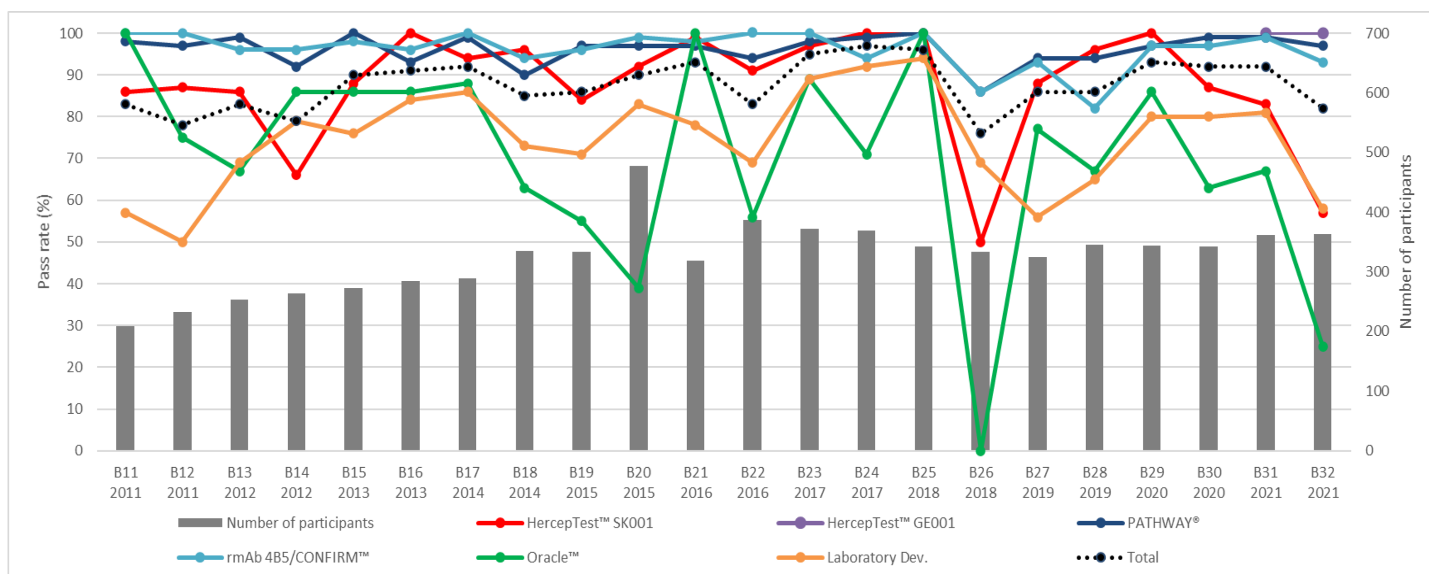


Table 1. **Assessment marks for IHC assays and antibodies run B32, HER2 IHC**

IVD approved HER2 assays	n	Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
PATHWAY® rmAb clone **4B5, 790-2991, (VRPS)[4]**	32	Ventana/Roche	23	8	1	-	97%	72%
PATHWAY® rmAb clone **4B5, 790-2991, (LMPS)[5]**	122	Ventana/Roche	84	29	6	3	93%	69%
rmAb clone **4B5, 790-4493, (VRPS)[4]**	17	Ventana/Roche	11	4	1	1	88%	65%
rmAb clone **4B5, 790-4493, (LMPS)[5]**	65	Ventana/Roche	44	17	2	2	94%	68%
HercepTest™, pAb **SK001, (VRPS)[4]**	14	Dako/Agilent	4	4	1	5	57%	29%
HercepTest™, pAb **SK001, (LMPS)[5]**	6	Dako/Agilent	-	2	3	1	33%	-
HercepTest™, rmAb **DG44 GE001, (VRPS)[4]**	16	Dako/Agilent	8	8	-	-	100%	50%
HercepTest™, rmAb **DG44 GE001, (LPMS)[5]**	2	Dako/Agilent	1	1	-	-	-	-
Oracle™ mAb clone **CB11, TA9145, (VRPS)[4]**	4	Leica Biosystems	-	1	1	2	-	-
Oracle™ mAb clone **CB11, TA9145, (LMPS)[5]**	4	Leica Biosystems	1	2	-	1	-	-
Antibodies[3] for laboratory developed HER2 assays, conc. antibody		Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
mAb clone **10A7**	1	Leica Biosystems	-	-	-	1	-	-
mAb clone **CB11**	1	Leica Biosystems	-	-	-	1	-	-
mAb clone **IHC012**	1	GenomeMe	-	1	-	-	-	-
rmAb clone **BP6020**	1	Bailing Biotechnology	-	1	-	-	-	-
rmAb clone **EP3**	1 1	Cell Marque Epitomics	1	1	-	-	-	-
rmAb clone **QR3**	1	Quartett	-	-	-	1		
rmAb clone **RM228**	1	Revmab Biosciences	-	-	-	1		
rmAb clone **SP3**	8 6 2 1 1 1	Thermo Fisher Scientific Cell Marque Zytomed DCS Histols/Diagmics Master Diagnostica	2	6	4	7	42%	11%
pAb, **A0485**	46	Dako/Agilent	17	13	4	12	65%	37%
Antibodies for laboratory developed HER2 assays, RTU		Vendor	Optimal	Good	Borderline	Poor	Suff.[1]	OR[2]
Ab clone **GR011, 8362-C010**	2	Sakura Finetek	-	-	1	1	-	-
rmAb clone **EP3 AN726**	1	BioGenex	-	1	-	-		
rmAb clone **EP3 ARMPD049R**	1	Diagnostic Biosystems	1					
rmAb clone **SP3, MAD-000308QD**	2	Master Diagnostica	1	1	-	-	-	-
rmAb clone **SP3, 237R-17**	2	Cell Marque	-	1	1	-	-	-
Total	363		198	101	25	39		
Proportion			54%	28%	7%	11%	82%	

1) Suff.; Proportion of sufficient stains (optimal or good).
2) OR; Proportion of optimal results.
3) mAb: mouse monoclonal antibody, rmAb: rabbit monoclonal antibody, pAb: polyclonal antibody.
4) VRPS; Vendor Recommended Protocol Settings – RTU system used in compliance to protocol settings and package insert.
5) LMPS; Laboratory Modified Protocol settings - RTU system used by modified protocol settings focusing on retrieval conditions, Ab incubation time, detection system and IHC platform.

Detailed Analysis
IVD approved assays

PATHWAY® rmAb clone **4B5** (790-2991, Ventana/Roche): In total, 107 of 154 (69%) protocols were assessed as optimal. Protocols with optimal results were typically based on Heat Induced Epitope Retrieval (HIER) in Cell Conditioning 1 (CC1) (efficient heating time 16-64 min.) on BenchMark XT, GX or Ultra, 12-32 min. incubation of the primary Ab and iView or UltraView as detection kit. Using these protocol settings, 118 of 124 (95%) laboratories produced a sufficient staining result (optimal or good).

rmAb clone **4B5** (790-4493, Ventana/Roche): In total, 55 of 82 (67%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in CC1 (efficient heating time 32-64 min.) on BenchMark XT, GT or Ultra, 12-32 min. incubation of the primary Ab and UltraView as detection system. Using these protocol settings, 63 of 67 (94%) laboratories produced a sufficient staining result.

HercepTest™ pAb (SK001, Dako/Agilent): In total, 4 of 20 (20%) protocols were assessed as optimal. Protocols with optimal results were typically based on HIER in HercepTest™ epitope retrieval solution at 97-99°C for 40-45 min. in a water bath or PT Link, 30 min. incubation of the primary Ab and SK001 Polymer as detection system. Using these protocol settings, 9 of 15 (60%) laboratories produced a sufficient staining result.

HercepTest™ rmAb clone **DG44** (GE001, Dako/Agilent): In total, 9 of 18 (50%) protocols were assessed as optimal. Protocols with optimal results were based on HIER in HercepTest™ epitope retrieval solution at 97°C for 30 min., 10 min. incubation of the primary Ab and GE001 Polymer as detection system. Using these protocol settings, 17 of 17 (100%) laboratories produced a sufficient staining result.

Table 2 summarizes the proportion of sufficient and optimal marks for the most commonly used IVD approved assays. The performance was evaluated both as "true" plug-and-play systems performed accordingly to the vendor recommendations and by laboratory modified systems changing basal protocol settings. <u>Only protocols performed on the specific IHC stainer device are included.</u>

Table 2. **Comparison of pass rates for vendor recommended and laboratory modified protocols**

CDx assay	**Vendor recommended protocol settings***		**Laboratory modified protocol settings****	
	Sufficient	Optimal	Sufficient	Optimal
Ventana BenchMark XT, GX, Ultra PATHWAY® rmAb 4B5, **790-2991**	31/32 (97%)	23/32 (72%)	111/119 (93%)	83/119 (70%)
Ventana BenchMark XT, GX, Ultra rmAb 4B5, **790-4493**	15/17 (88%)	11/17 (65%)	58/62 (94%)	41/62 (66%)
Dako Autostainer Link 48+ HercepTest™ pAb **SK001**	8/14 (57%)	4/14 (29%)	2/4	0/4
Dako Omnis HercepTest™ rmAb DG44, **GE001**	16/16 (100%)	8/16 (50%)	2/2	1/2
Leica Bond MAX, III Oracle™ mAb CB11, **TA9145**	1/4	0/4	3/4	1/4

* Protocol settings recommended by vendor – Retrieval method & conditions, Ab incubation times, detection kit, IHC stainer/equipment.
** Modifications included: retrieval method, retrieval duration, retrieval reagents, Ab incubation time and detection kit. Only protocols performed on the specified vendor IHC stainer were included.

Concentrated antibodies for laboratory developed (LD) assays

pAb, **A0485**: 17 of 46 (56%) protocols were assessed as optimal. Optimal protocols were based on HIER using either Target Retrieval Solution (TRS) low pH (Dako/Agilent) (10/26*), TRS High pH (Dako/Agilent) (6/13) or CC1 (Ventana/Roche) (1/4). The Ab was diluted in the range of 1:100-1,600 depending on the level of the total technical sensitivity of the protocol employed. Using these protocol settings, 28 of 43 (65%) laboratories produced a sufficient staining result.

* (number of optimal results/number of laboratories using this HIER buffer)

Table 3 summarizes the overall proportion of optimal staining results when using the most frequently used concentrated Ab on the most commonly used IHC stainer platforms.

Table 3. **Optimal results for HER2 for the most commonly used antibody as concentrate on the four main IHC systems***

Concentrated antibodies	**Dako/Agilent Autostainer**		**Dako/Agilent Omnis**		**Ventana/Roche BenchMark GX / XT / Ultra**		**Leica Biosystems Bond III / Max**	
	TRS High pH	TRS Low pH	TRS High pH	TRS Low pH	CC1 pH 8.5	CC2 pH 6.0	BERS2 pH 9.0	BERS1 pH 6.0
pAb clone **A0485**	2/7** (29%)	4/10 (40%)	4/6 (67%)	6/16 (38%)	1/4	-	0/3	-

* Antibody concentration applied as listed above, HIER buffers and detection kits used as provided by the vendors of the respective platforms.
** (number of optimal results/number of laboratories using this buffer)

Comments

In this NordiQC assessment B32 for HER2, an overall pass rate of 82% was observed which was slightly reduced to the level seen in the latest assessment runs, B29-B31.
The insufficient results were primarily characterized by a too weak / false negative staining reaction being observed in 70% (45 of 64 results). Virtually all laboratories were able to demonstrate the expected HER2 3+ staining reaction in the breast carcinomas, tissue cores no. 1 and 4, with high level gene amplification, whereas too weak / false negative staining results were particularly and most critically observed as a 0/1+ IHC staining reaction in the HER2 gene amplified breast carcinoma, tissue core no. 5. This tumour was categorized as IHC 2+ in the NordiQC reference laboratories using the two FDA/CE-IVD HER2 IHC assays: PATHWAY® (Ventana/Roche) and HercepTest™ (Dako/Agilent) and showed HER2 gene amplification (ratio 3.6) by FISH.
In the remaining insufficient results, these were characterized by e.g. a poor signal-to-noise ratio, impaired morphology and/or excessive cytoplasmic staining reaction compromising the interpretation of the specific HER2 membranous reaction.

76% of the participants (n=274) used FDA/CE-IVD approved companion diagnostic (CDx) HER2 IHC assays as PATHWAY® (Ventana/Roche), HercepTest™ (Dako/Agilent) and Oracle™ (Leica Biosystems) on the specified stainer with predictive claim for HER2 status in breast cancer. 2% (n=8) of the participants used an approved assay on another platform than specified by the vendor, while the remaining 22% (n=81) used a laboratory developed test (LDT) based on a concentrated primary Ab or a RTU format without a predictive claim. This segmentation has been relatively consistent in the last assessment runs.

The Ventana/Roche PATHWAY® HER2 IHC assays 790-2991 and 790-4493 were used by 65% of all participants (n=236). When applying the assays on the intended platform, BenchMark, an overall pass rate of 93% was observed and 69% were optimal. In both the previous and this assessment, the pass rates and proportion of optimal results for laboratories using these two IHC assays as "plug-and-play" and strictly compliant to the recommended protocol settings or using modified protocols were fully comparable as seen in Table 1 and 2. Despite this observation, it is still highly recommended to use the assays strictly in concordance to the instructions and guidelines provided by the vendor, as e.g. in run B28 it was shown that both the pass rate and proportion of optimal results were reduced for laboratories modifying the protocols. More data can be found at; https://www.nordiqc.org/downloads/assessments/123_11.pdf

Similar to runs B30 and B31, it was observed that an increased number of participants used OptiView or UltraView with amplification for the HER2 IHC assays 790-2991 and 790-4493, substituting iView or UltraView as recommended by Ventana/Roche. In this run 10% of the laboratories used one of the two HER2 CDx assays in combination with either OptiView or UltraView with amplification, which was the same level seen in run B28. In run B28 this modification frequently induced an insufficient result characterized by a false positive HER2 reaction in a 2+ HER2 gene unamplified breast carcinoma. This underlines that modifications of CDx assays should be meticulously validated by the end-users on a large cohort of breast carcinomas (e.g. n=100). This has been addressed by ASCO/CAP in both the 2013 guidelines for HER2 testing and the 2020 guidelines for ER/PR testing and in particular in detail in the publication by Torlakovic et al; *"Evolution of Quality Assurance for Clinical Immunohistochemistry in the Era of Precision Medicine Part 3: Technical Validation of Immunohistochemistry", AIMM 2017;25:151–159*

The recently launched Dako/Agilent HercepTest™ CDx assay for Dako Omnis based on the rmAb clone DG44 was used by 5% (n=18) of all participants and provided a pass rate of 100%, 50% optimal results. As seen in Table 2, the majority of all laboratories used the assay by vendor recommended protocol settings. The proportion of optimal results was reduced compared to the level seen for the Ventana/Roche

CDx assays based on rmAb clone 4B5 and mainly caused by a slightly increased cytoplasmic staining reaction complicating the interpretation of the coexisting membranous reaction.

The Dako/Agilent HercepTest™ CDx assay SK001 for Dako Autostainer Link 48 provided a low pass rate of 57% and only 29% optimal results This low level pass rate was obtained despite the IHC CDx assay was used in concordance with the recommended protocol settings from Dako/Agilent. This observation and inferior performance of the assay has been seen in the latest 3 assessment runs for HER2 IHC as shown in Graph 1. At present no plausible cause for this decline can be identified and laboratories are encouraged to verify the performance of the assay and total system including the Autostainer.

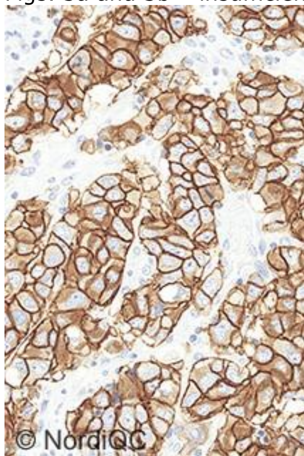
In this HER2 IHC assessment, 22% of the participants used LDTs based on concentrated Ab formats or generic RTU Abs without intended use or predictive claim for HER2 demonstration in breast carcinoma to guide decision with treatment with Herceptin or similar drugs. Overall, the LDTs provided a pass rate of 58% (47 of 81) and 27% optimal (22 of 81).
The pAb A0485 from Dako/Agilent was the most widely applied Ab within a LDT being used by 13% of the participants and when applied with optimal protocol settings as described above, a pass rate of 65% was obtained.
The rmAb clone SP3 as concentrate was used by 5% of the participants and similar to the data seen in the latests assessment runs found less successful. In this assessment run B32, a pass rate of 42% was seen.

The proportion of laboratories using the FDA-/CE-IVD approved HER2 IHC assays and LDTs is very consistent. In this run, 22% of the participants (n=81) used LDTs compared to 23-31% in the latest assessments.

Scoring consensus B32
Laboratories were requested to submit scores (0, 1+, 2+ or 3+) on the NordiQC homepage of their own HER2 stained slides. This was done by 83% (300 of 363) of the participants returning slides.
For 245 of the 300 (82%) responding participants, scores for all the tissues in the multi-tissue sections were in concordance with the NordiQC assessor group using the ASCO/CAP 2018 scoring guidelines. This was slightly improved to the level of of 75% and 77% observed in runs B30 and B31, respectively.
Among laboratories with sufficient staining, 84% (210 of 249) of the scoring read-outs were in agreement with the NordiQC assessors. Disagreement was primarily related to the scoring of the HER2 status in breast carcinoma, tissue core no. 4. This was characterized as 3+ both by the NordiQC reference standard methods and by the vast majority of all participants. The membranes of neoplastic cells in the tumour, however were less intense compared to the breast carcinoma, tissue core no. 1, being very intense and simultaneously tissue core no. 4 in areas showed a more extensive cytoplasmic staining reaction. However, both tumours should be scored as 3+, accordingly to the ASCO/CAP 2018 scoring guidelines.
Among participants with insufficient staining results, 69% were in consensus with the NordiQC assessor group (35 of 51) and this was an improvement compared to the level of 37% in run B31. For this group the disagreement primarily was related to the scoring of the breast carcinoma, tissue core no. 5. The results submitted to NordiQC was scored as 1+ by NordiQC assessor team and typically as 2+ by the participant. The NordiQC assessment was primarily based on strict adherence to the ASCO/CAP guidelines but also to the level expected and characterized by the two HER2 IHC reference standard methods.

Conclusion
The FDA-/CE-IVD approved HER2 IHC assays **PATHWAY®/4B5** 790-2991/790-4493 from Ventana/Roche and **HercepTest™**, GE001 Dako/Agilent were in this assessment the most accurate and successful assays for the semi-quantitative IHC determination of HER2 protein expression in breast carcinoma.
Laboratory developed assays based on concentrated formats and RTU formats without a predictive claim provided a lower pass rate and reduced proportion of optimal results.
Inclusion of 2+ tumours with and without HER2 gene amplification in the control material for both EQA and internal quality control seems to be essential to evaluate accuracy, precision and reproducibility of the HER2 IHC assays used by laboratories.

Figs. 1a and 1b – optimal staining results, same protocol
Figs. 2a and 2b – insufficient staining results - false negative, same protocol
Figs. 3a and 3b – insufficient staining results – poor signal-to-noise ratio, same protocol



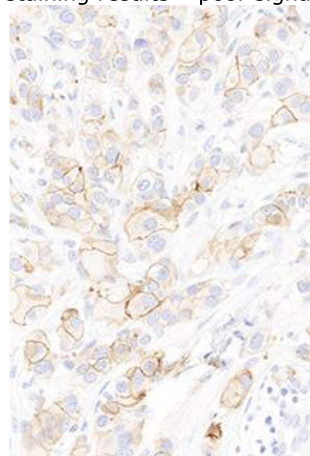
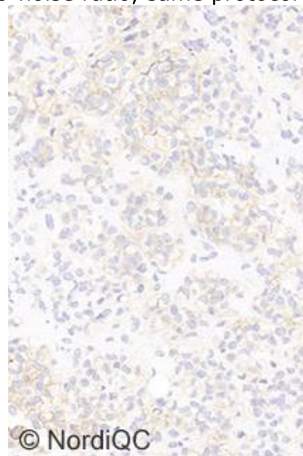


Fig. 1a.
Left: Optimal staining result for HER2 of the breast carcinoma no. 4 with a ratio of HER2 / chr17 of 6.7.
> 10% of the neoplastic cells show a strong and complete membranous staining reaction corresponding to 3+.
Right: Optimal staining result for HER2 of the breast carcinoma no. 5 with a ratio of HER2 / chr17 of 3.6.
> 10% of the neoplastic cells show a weak to moderate and complete membranous staining reaction corresponding to 2+.



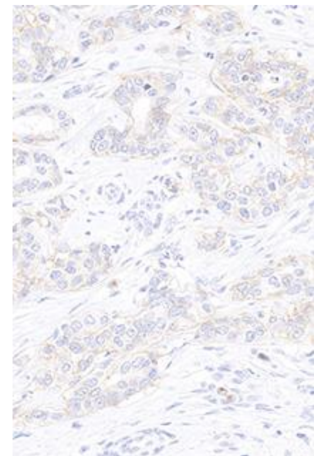
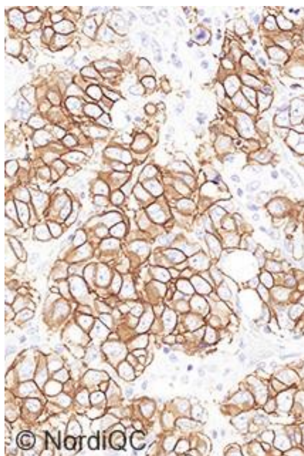


Fig. 1b.
Left: Optimal staining result for HER2 of the breast carcinoma no. 2 with a ratio of HER2 / chr17 of 1.0.
> 10% of the neoplastic cells show a weak complete membranous staining reaction corresponding to 2+.
Right: Optimal staining result for HER2 of the breast carcinoma no. 3 with a HER2 / chr17 ratio of 1.1.
> 10% of the neoplastic cells show a faint, partial membranous staining reaction corresponding to 1+.



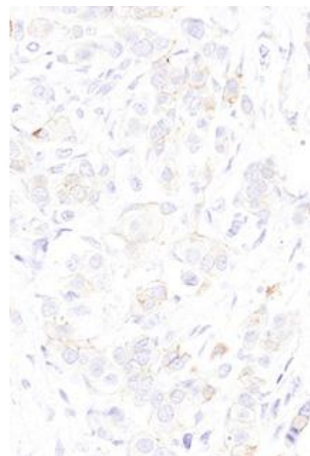
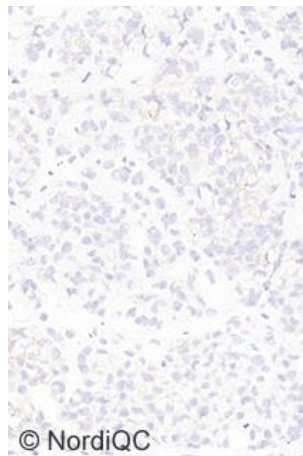


Fig. 2a.
Left: Staining result for HER2 of the breast carcinoma no. 4 with a ratio of HER2 / chr17 of 6.7.
> 10% of the neoplastic cells show a strong complete membranous staining reaction corresponding to 3+.
Right: **Insufficient and false negative staining result** for HER2 of the breast carcinoma no. 5 with a ratio of HER2 / chr17 of 3.6
> 10% of the neoplastic cells show a weak to moderate, but incomplete membranous staining reaction corresponding to 1+ (the core was scored as 1+ both by the participant and NordiQC).



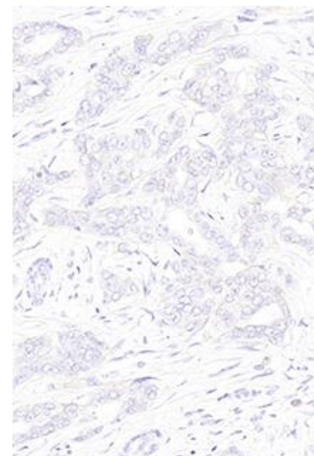
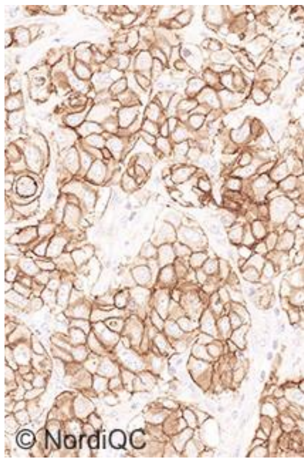


Fig. 2b.
Left: Staining result for HER2 of the breast carcinoma no. 2 with a ratio of HER2 / chr17 of 1.0.
< 10% of the neoplastic cells show a weak partial membranous staining reaction corresponding to 0.
Right: Staining result for HER2 of the breast carcinoma no. 3 with a HER2 / chr17 ratio of 1.1.
No staining reaction is seen corresponding to 0.



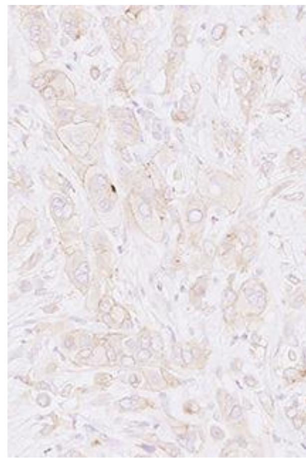
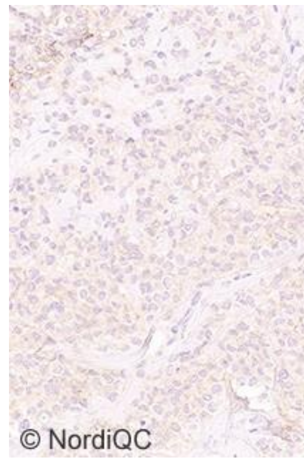


Fig. 3a.
Left: Staining result for HER2 of the breast carcinoma no. 4 with a ratio of HER2 / chr17 of 6.7.
> 10% of the neoplastic cells show an intense and complete membranous staining reaction corresponding to 3+.
Right: **Insufficient staining result** for HER2 of the breast carcinoma no. 5 with a ratio of HER2 / chr17 of 3.6. The excessive cytoplasmic staining reaction and reduced membranous staining reaction comprises the read-out. The tumour was scored as 1+ by the participant.



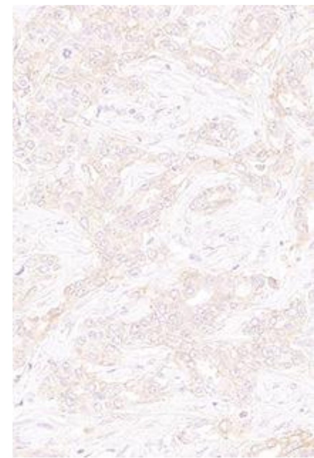


Fig. 3b.
Left: Staining result for HER2 of the breast carcinoma no. 2 with a ratio of HER2 / chr17 of 1.0.
The excessive cytoplasmic staining reaction and weak membranous reaction in the neoplastic cells compromises the read-out.
Right: Insufficient staining result for HER2 of the breast carcinoma no. 3 with a HER2 / chr17 ratio of 1.1.
As for the other cores a poor signal-to-noise is observed hampering the read-out.

SN/LE 06.12.2021